

## International Journal of Remote Sensing

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/tres20>

### Automated cloud and shadow detection and filling using two-date Landsat imagery in the USA

Suming Jin<sup>a</sup>, Collin Homer<sup>b</sup>, Limin Yang<sup>c</sup>, George Xian<sup>a</sup>, Joyce Fry<sup>b</sup>, Patrick Danielson<sup>c</sup> & Philip A. Townsend<sup>d</sup>

<sup>a</sup> ASRC Research and Technology Solutions, Contractor to the US Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD, 57198, USA

<sup>b</sup> US Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD, 57198, USA

<sup>c</sup> Stinger Ghaffarian Technologies, Contractor to the US Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD, 57198, USA

<sup>d</sup> Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, Madison, WI, 53706, USA

Version of record first published: 16 Oct 2012.

To cite this article: Suming Jin, Collin Homer, Limin Yang, George Xian, Joyce Fry, Patrick Danielson & Philip A. Townsend (2013): Automated cloud and shadow detection and filling using two-date Landsat imagery in the USA, International Journal of Remote Sensing, 34:5, 1540-1560

To link to this article: <http://dx.doi.org/10.1080/01431161.2012.720045>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any

instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## Automated cloud and shadow detection and filling using two-date Landsat imagery in the USA

Suming Jin<sup>a\*</sup>, Collin Homer<sup>b</sup>, Limin Yang<sup>c</sup>, George Xian<sup>a</sup>, Joyce Fry<sup>b</sup>,  
Patrick Danielson<sup>c</sup>, and Philip A. Townsend<sup>d</sup>

<sup>a</sup>ASRC Research and Technology Solutions, Contractor to the US Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD 57198, USA; <sup>b</sup>US Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD 57198, USA; <sup>c</sup>Stinger Ghaffarian Technologies, Contractor to the US Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD 57198, USA; <sup>d</sup>Department of Forest and Wildlife Ecology, University of Wisconsin–Madison, Madison, WI 53706, USA

(Received 3 August 2011; accepted 10 January 2012)

A simple, efficient, and practical approach for detecting cloud and shadow areas in satellite imagery and restoring them with clean pixel values has been developed. Cloud and shadow areas are detected using spectral information from the blue, shortwave infrared, and thermal infrared bands of Landsat Thematic Mapper or Enhanced Thematic Mapper Plus imagery from two dates (a target image and a reference image). These detected cloud and shadow areas are further refined using an integration process and a false shadow removal process according to the geometric relationship between cloud and shadow. Cloud and shadow filling is based on the concept of the Spectral Similarity Group (SSG), which uses the reference image to find similar alternative pixels in the target image to serve as replacement values for restored areas. Pixels are considered to belong to one SSG if the pixel values from Landsat bands 3, 4, and 5 in the reference image are within the same spectral ranges. This new approach was applied to five Landsat path/rows across different landscapes and seasons with various types of cloud patterns. Results show that almost all of the clouds were captured with minimal commission errors, and shadows were detected reasonably well. Among five test scenes, the lowest producer's accuracy of cloud detection was 93.9% and the lowest user's accuracy was 89%. The overall cloud and shadow detection accuracy ranged from 83.6% to 99.3%. The pixel-filling approach resulted in a new cloud-free image that appears seamless and spatially continuous despite differences in phenology between the target and reference images. Our methods offer a straightforward and robust approach for preparing images for the new 2011 National Land Cover Database production.

### 1. Introduction

Landsat is one of the most widely used satellite data sources for local, regional, and global applications because of its medium spatial resolution, multispectral bands, and long record of historical data (Vogelmann et al. 2001; Homer et al. 2004; Goward et al. 2006; Williams, Goward, and Arvidson 2006). With its no-cost policy, application of Landsat data will further increase (Woodcock et al. 2008). However, clouds and cloud shadows often impede

---

\*Corresponding author. Email: [sjin@usgs.gov](mailto:sjin@usgs.gov)

many applications such as land-cover classification and change detection because of the spectral variance they introduce (Ju and Roy 2008). In the USA, one primary use of Landsat data has been the production of the National Land Cover Database (NLCD), with *circa* 1992, 2001, 2006, and soon 2011 epochs produced under the auspices of the Multi-Resolution Land Characteristics (MRLC) Consortium (<http://www.mrlc.gov>). The MRLC data pool now includes over 16,000 cloud-free or nearly cloud-free Landsat 5 and Landsat 7 images, which have been consistently preprocessed and are freely available for download to the public. Historically, MRLC has included few cloudy scenes; the clouds and shadows in these scenes were removed through hand digitizing. However, the new 2011 NLCD is designed to use multi-temporal and multi-date Landsat imagery for change detection and land-cover classification; therefore, an automated and widely applicable cloud and shadow detection and removal method is needed to increase the number of Landsat images available for such applications.

Few methods of cloud detection are operational at a regional scale or beyond. The automated cloud-cover assessment (ACCA) algorithm has been operationally used primarily to estimate Landsat cloud-cover percentage for guiding scene selection and optimizing scene acquisition (Irish 2000; Irish et al. 2006). The ACCA is composed of 26 specific decisions, or filters, which make full use of spectral information from band 2 to band 6 and their derivative indices. 'ACCA clouds' are defined as optically thick or nearly opaque. Roy et al. (2010) developed cloud masks for web-enabled Landsat data (Landsat Enhanced Thematic Mapper Plus (ETM+) composited mosaics of the conterminous USA) using a classification tree in addition to ACCA. They developed statistical classification rules using the optical and thermal wavelength values from training pixels. Both ACCA and classification tree methods fully explored the capability of spectral values, but not the temporal information to achieve their goals in developing cloud masks. The Multi-Temporal Cloud Detection (MTCD) method has been developed in the framework of the preparation of VENμS and SENTINEL-2 Level 2 processors (Hagolle et al. 2010). The MTCD method detects sudden increases of reflectance in the blue wavelength on a pixel-by-pixel basis using multi-temporal observations and tests the linear correlation of pixel neighbourhoods taken from pairs of images acquired successively. Hagolle et al. (2010) concluded that the MTCD method provides a better discrimination of clouded and unclouded pixels than the ACCA methods. However, the MTCD method requires high revisit frequency satellite data and chronological processing of the data. One common limitation of these methods is that shadows from clouds are not assessed.

Many other methods have been developed for both cloud and shadow detection at local and/or regional scales. Ho and Cai (1996) used supervised classification to map cloud and shadow areas, then eliminated false clouds and shadows using geometric constraints between clouds and their shadows. They summarized that it is normal that many ground objects may also fall into cloud and shadow classes through supervised and/or unsupervised classification. They further observed that geometric constraints were very effective in eliminating false cloud and shadow pairs. Wang et al. (1999) detected clouds using the brightness of the target image and brightness difference between the target image and a reference image. They detected shadows by means of a wavelet transform because shadows smooth the brightness changes of the ground. A reference image with no overlapping clouds and shadows with the target image was selected. Martinuzzi, Gould, and Ramos González (2007) created cloud masks using Landsat ETM+ band 1 (blue) and thermal band 6. They then created a potential shadow mask by displacement of the cloud mask based on the horizontal shift with a 10-pixel buffer added. Within the shadow mask, the brightness values in band 4 were further used to differentiate cloud-shadow areas from nonshadowed areas. Choi and Bindschadler (2004) designed an algorithm to detect clouds

on ice sheets using the normalized difference snow index (NDSI) and a shadow matching technique according to the knowledge of the solar illumination azimuth. NDSI is the normalized difference between band 2 (green) and band 5 (mid-infrared (MIR)). The matching procedure works with the sets of edges of possible clouds, the possible cloud shadows, and the water regions. Hégarat-Masclé and André (2009) inter-calibrated bands 2 and 5 to detect clouds and MIR to detect shadows. Relationships between clouds and their corresponding shadows, in particular their similar shape, area, and relative locations, were formalized using Markov random fields. The relationship was then used for cloud and shadow mutual validation. Huang et al. (2010) developed an algorithm that used clear view forest pixels as a reference to define cloud boundaries in a spectral-temperature space. Furthermore, shadow locations were predicted according to cloud height estimates and sun illumination geometry, and actual shadow pixels were identified by searching the darkest pixels surrounding the predicted shadow locations. Their algorithm was particularly designed for forest change analysis using Landsat images.

Besides cloud and shadow automatic detection, various techniques, which we grouped into four general categories, have been used to fill gaps caused by cloud and shadow or in scan line corrector (SLC)-off images.

- (1) *Compositing*. This method replaced contaminated pixels with cloud-free pixels from a reference image (Martinuzzi, Gould, and Ramos González 2007; Roy et al. 2010). Roy (2000) summarized effective compositing criteria, which include the maximum normalized difference vegetation index (NDVI), maximum brightness temperature, maximum apparent surface temperature, maximum difference in red and near-infrared (NIR) reflectance, minimum scan angle, and combinations of these. Because the compositing method depends solely on the reference image, the quality of the filled value will be heavily affected by any difference between the reference and target images, which can be caused by many things such as phenology, preprocessing, atmospheric condition, and disturbance.
- (2) *Regression*. This method uses the regression relationship and/or histogram matching between the target image and a reference image to generate a cloud-free image (Helmer and Ruefenacht 2005). Earth Satellite Corporation has developed software and a method that uses classification and regression tree technology to eliminate cloud and shadow areas (Herold, Cunningham, and Wylie 2003). The regression method tends to have difficulties with heterogeneous landscapes and linear features because of the detail loss that occurs with regression tree prediction.
- (3) *Data fusion*. Wang et al. (1999) adopted wavelet transformation to incorporate complementary information into the composite from multi-temporal images. Roy et al. (2008) developed a semi-physical fusion approach that used the Moderate Resolution Imaging Spectroradiometer (MODIS) Bidirectional Reflectance Distribution Function (BRDF)/Albedo land surface characterization product and Landsat ETM+ data to predict ETM+ reflectance on the same date. This method is applicable, provided that the wavelengths considered are similar to those available from MODIS. The authors mentioned that it is not trivial to do data co-registration and computation of Landsat satellite viewing and solar geometry.
- (4) *Neighbourhood pixel interpolation*. This approach uses cloud-free pixels from the image itself to predict the values for contaminated areas. Approaches can vary and include using focal mean, nearest neighbour (Chen et al. 2011), or geostatistics methods (Zhang, Li, and Travis 2007) to fill gaps. This method can be computationally intensive and will predict poorly at the centre of large cloud/shadow patches because of its reliance on neighbourhood pixels.

The above cloud and shadow detection methods used brightness, temperature, temporal information, geometric relationships between cloud and shadow, or some combination of these kinds of information. However, little research has been done to develop a comprehensive and widely applicable module flexibly integrating different strategies to produce cloud and shadow masks. Each of the above gap-filling methods has its own set of strengths and weaknesses. Some of these methods are heavily dependent on the reference image, whereas others do not depend on the reference image at all. Some methods are computationally intensive. Some methods have difficulties with heterogenous landscapes and linear features.

In this study, we propose a simple and practical method that produces cloud and shadow masks by flexibly using spectral (blue, NIR, MIR, and thermal bands), temporal information from two-date imagery, and spatial relationship (size, distance, and direction) between cloud and corresponding shadow. We also propose an efficient, practical, and balanced method, which uses information from two-date images for filling cloud- and shadow-masked areas. The image for which we intend to detect clouds and shadows is referred to as the target image, and the image that is used for assistance in detecting and filling is referred to as the reference image.

2. Methodology

2.1. Data sets

We selected five Landsat path/rows (Table 1, Figure 1) representing different landscapes, seasons, and cloud patterns to demonstrate the effectiveness of our methods. We used two images per path/row: one target scene and one reference scene. The reference images should have no overlapping clouds with the target images. There is no requirement for a close match in phenology and image acquisition time between the target and reference scenes, even though it is preferable to select the two scenes acquired from the same season and within a short temporal gap if they are available. Landsat scene p43r32 (Figure 1(a)) covers the Black Rock Desert and Sierra Nevada terrain across Nevada and California. The main land-cover types are shrub, barren land, and forest. Clouds are mainly located on the upper part of the scene, including the desert. Landsat scene p42r28 (Figure 1(b)) is mainly composed of forests and agricultural lands. Clouds in the target scene are on the forest areas at the upper right corner of the scene. Landsat scene p33r33 (Figure 1(c)) is mainly composed of agriculture, forest lands, and urban land covers. Clouds are on the agricultural lands at the lower right corner of the target scene and cloud sizes vary from a few pixels

Table 1. Landsat data sets for cloud and shadow detection and filling.

Path/row	Target image	Reference image	Main land cover
p43r32	20 June 2010 (lt5)	11 July 2006 (lt5)	Shrub, barren land, evergreen forest
p42r28	11 September 1999 (le7)	28 August 2000 (le7)	Evergreen forest, cultivated crops, grassland
p33r33	16 October 2000 (le7)	14 September 2000 (le7)	Grassland, evergreen forest, Developed
p18r34	5 March 2000 (lt5)	30 April 2000 (le7)	Deciduous forest, pasture hay, barren land
p12r30	16 June 2006 (lt5)	29 April 2006 (lt5)	Evergreen forest, deciduous forest, water

Note: lt5, Landsat 5; le7, Landsat 7.



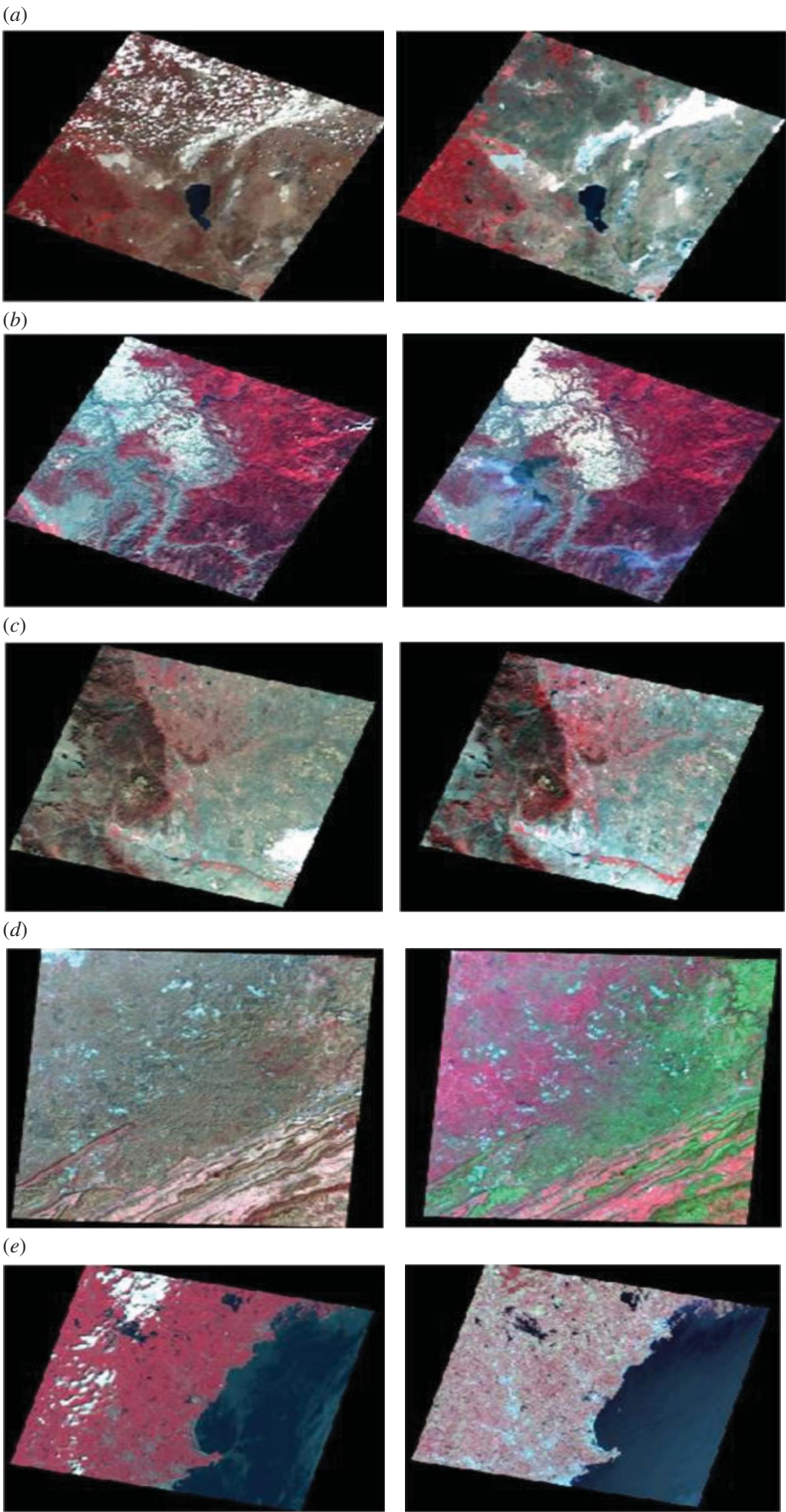


Figure 1. Target (left) and reference (right) images (displayed as bands 4, 3, and 2) for five Landsat path/rows: (a) p43r32, (b) p42r28, (c) p33r33, (d) p18r34, and (e) p12r30.

to hundreds of thousands of pixels. Landsat scene p18r34 (Figure 1(d)) covers mountainous terrain across Virginia, West Virginia, and Kentucky. Mining has been active there for centuries. Clouds are distributed widely in the target scene and some clouds are thin but nearly opaque. Landsat scene p12r30 (Figure 1(e)) is in a relatively flat area and mainly composed of forests and water. Clouds vary in size and type. Both p18r34 (Figure 1(d)) and p12r30 (Figure 1(e)) have very different phenology between their target and reference images.

All of the images were preprocessed following the MRLC 2001 protocol (Chander et al. 2009). Image pairs were geometrically corrected and converted to top-of-atmosphere reflectance, and their thermal bands were converted to at-sensor brightness temperature. The top-of-atmosphere reflectance values for all six reflective bands (bands 1–5 and 7) are multiplied by 400 to produce 8-bit data. The at-sensor brightness temperature is also rescaled by subtracting 240, then multiplying by 3 to produce 8-bit data. These images are all in rescaled 8-bit format and resampled to have a spatial resolution of 30 m.

## 2.2. Cloud and shadow detection

The cloud and shadow detection model consists of three major functional parts (Figure 2). The first part produces two (relaxed and restricted) cloud and shadow masks. The second part integrates both relaxed and restricted masks to use compensating features from both masks. The third part employs the geometric relationship between cloud and shadow (distance, size, and direction) to remove false shadows on the integrated cloud and shadow masks to produce a final constrained cloud and shadow mask for the target image.

### 2.2.1. Initial relaxed and restricted cloud and shadow masks

We incorporate Landsat blue and thermal bands from two-date images as the primary candidates to detect clouds. The first assumption is that clouds are normally brighter, especially

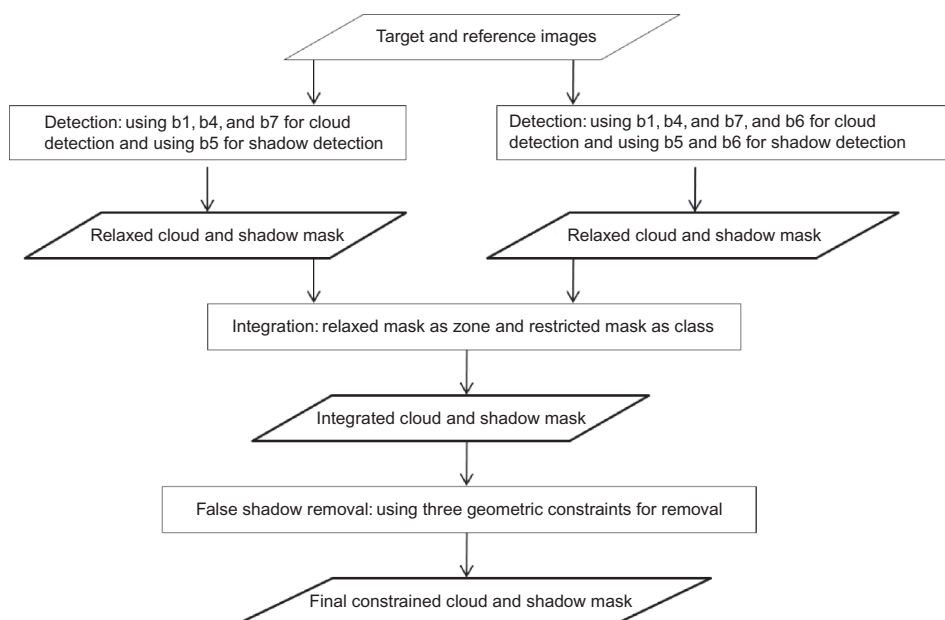


Figure 2. Flow chart of cloud and shadow mask development.



in the blue band, and are usually colder than the land surface they obscure (Bréon and Colzy 1999; Irish 2000). The second assumption is that the other bright land surfaces are relatively stable compared to the quick variations of the reflectance of pixels affected by clouds (Hagolle et al. 2010). The relaxed cloud class is detected using the blue band of the target image and the blue-band difference image derived from the target and reference images. In addition, band 4 (NIR) and band 7 (shortwave infrared (SWIR)) of the target image are used to assist in removing some commission errors by eliminating pixels with relatively low spectral values in bands 4 and 7. Thermal bands from target and reference images are used to further refine the relaxed cloud class to produce a restricted cloud class.

Shadows are detected using band 5 (MIR) and the thermal bands from the two-date images in a similar way to clouds. The basic assumption is that shadows are normally darker and colder than the land surface they obscure (Irish 2000; Hégarat-Masclé and André 2009). The relaxed shadow class is derived only from two MIR bands of the target and reference images. The restricted shadow class is obtained after integrating information from the thermal bands as well as the MIR bands. The particular condition statements for the cloud and shadow detection are shown in Figure 3. Threshold values are relative measures of spectral deviation from the mean value of a particular layer. Final outputs from the first functional part of the model include one relaxed cloud and shadow mask and one restricted cloud and shadow mask.

### 2.2.2. Cloud and shadow mask integration

The second functional part of the model integrates the two cloud and shadow masks (relaxed and restricted) into one mask (Figure 2). In some cases, neither cloud nor shadow mask from the first model is satisfactory, i.e. the variability of the land surface and the wide reflectance and temperature profiles of the clouds cause too many clouds and shadows (large commission errors) in the relaxed mask and too few clouds and shadows (large omission errors) in the restricted mask. The process of integration is conceptually similar to a multi-scale object-oriented classification. The integration idea views each patch of cloud or shadow in the relaxed mask as a zone or an object and in the restricted cloud and shadow mask as a class. If a minimum of five cloud or shadow pixels from the restricted mask in a

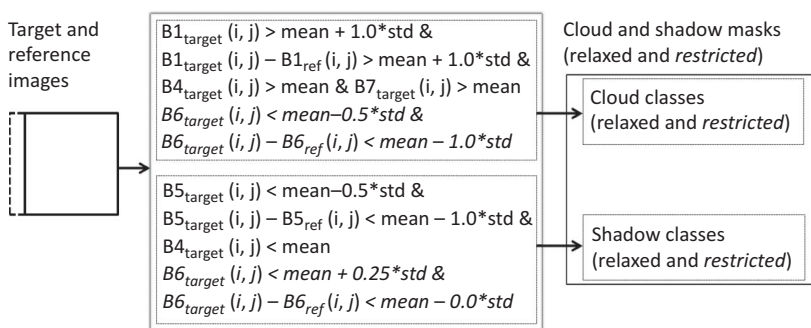


Figure 3. Relaxed and restricted cloud and shadow mask model conditions.  $B1_{\text{target}}(i, j)$  means band 1 pixel value of the location  $(i, j)$  in the target image.  $B1_{\text{ref}}(i, j)$  means band 1 pixel value of the location  $(i, j)$  in the reference image.  $B4_{\text{target}}(i, j)$  means band 4 pixel value of the location  $(i, j)$  in the target image.

zone or object exists, the entire zone or object (at least 10-pixel size) from the relaxed mask will be classified as cloud or shadow in the integrated mask; otherwise, the entire zone or object will disappear in the integrated mask.

### 2.2.3. False shadow removal

The third part of the model is designed to conservatively remove false shadows using three geometric constraints (Figure 2). The particular conditions of removing false shadows are listed below.

- (1) *Closeness*. A cloud and its shadow should not be far apart (Ho and Cai 1996). If no pixel of a shadow patch within the 100-pixel buffer zone of the cloud exists, the shadow patch is removed. We observed that the shortest distance between a cloud and its shadow is rarely beyond 100 pixels across the USA and acquisition seasons. The 100-pixel distance is the closest distance between a cloud and shadow patch, not the distance between a cloud pixel and its corresponding shadow.
- (2) *Size*. Shadow size should not be much bigger than its cloud (Ho and Cai 1996). However, a shadow patch may sometimes connect to areas that look like shadows, but are not actually shadows. Therefore, we conservatively remove any shadow whose size is equal to or larger than twice the corresponding cloud patch size.
- (3) *Direction*. Clouds and their shadows are related by their relative locations depending on the image acquisition time and the sun location (Hégarat-Masclé and André 2009). In the USA, all of the shadows are located northwest of the clouds because of the morning acquisition time. Therefore, we removed all shadows not located in that northwest direction.

### 2.3. Cloud and shadow filling

Pixels having the same spectral values are likely to be of the same land-cover type, and those pixels will likely have similar spectral values on a different date if no significant environmental and land-cover changes have taken place (Chen et al. 2011). Pixels with similar spectral values are referred to as a spectral similarity group (SSG). Our method for filling clouds and shadows is based on the concept of SSG. Landsat bands 3, 4, and 5 of the reference image are used to define a specific SSG for each contaminated pixel from the cloud and shadow mask of the target image. Pixels belong to a specific SSG when values from Landsat bands 3, 4, and 5 are within  $\pm 1$  deviation of the same band values.

We followed three steps (Figure 4) to fill a single cloud or shadow pixel. (1) The corresponding pixel in the reference image was identified for a cloud or shadow pixel within the target image, and the bands 3, 4, and 5 values of the pixel (referred as b3c, b4c, and b5c) were compared with every other pixel in the reference image. All other pixels form an SSG for the contaminated pixel if they have bands 3, 4, and 5 values within the ranges  $b3c \pm 1$ ,  $b4c \pm 1$ , and  $b5c \pm 1$ , respectively. (2) The locations of SSG pixels from the reference image are transferred to the target image, and the mean values of those pixels from the target image are calculated. (3) The mean values are used to replace the contaminated pixel values for the cloud or shadow pixel. The process is repeated for each pixel from the cloud and shadow mask. Because cloud size and location varies, the entire reference image is searched for SSG. After the process, very few isolated pixels are left unfilled, and the local

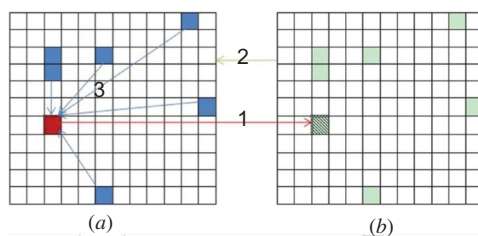


Figure 4. Cloud-filling diagram: (a) simulated target image; (b) simulated reference image of the same area. The red and dashed cell is a contaminated pixel, and the green and dashed cell is the corresponding cell on the reference image.

Notes: Green cells belong to one SSG identified using the reference image and are cloud free. Blue cells are the location-projected pixels from green cells and are also cloud free. Mean values from the blue cells are used to replace the values of the red cell.

mean method is applied to those pixels to complete the filled image. All of these processes are integrated in one single model.

Before filling, we buffered five pixels around the clouds and shadows on the constrained mask because the cloud and shadow mask is not perfect at a single pixel level, especially regarding the cloud edges.

## 2.4. Accuracy assessment

An accuracy assessment was conducted using a procedure based on a statistical sampling design. Specifically, a stratified random sampling procedure was applied, in which three strata (classes) were formed, with each containing cloud, shadow, and cloud-free pixels. A number of random points was drawn from each stratum for assessing the accuracy of the method.

For three Landsat images (p43r32, p42r28, and p33r33), we stratified and randomly created 300 points with a minimum of 50 points for each class on the final cloud and shadow mask, i.e. the constrained mask. For p18r34 and p12r30, we increased the total points to 500 and the minimum points to 100 because these scenes had clouds and shadows widely distributed throughout the image. Visual inspection of the Landsat images and historical aerial images from Google Earth were used to collect reference data for each selected sampling pixel, and a reference label was assigned to one of the three strata/classes (cloud, shadow, and cloud-free pixels).

The outcome of the sampling and interpretation process was a complete reference data set, which was used to assess the accuracy of each of the four output files obtained from the cloud and cloud shadow detection algorithms. The four output files of cloud and shadow masks are: constrained, integrated, relaxed, and restricted, as defined in Section 2.2. For each output file, an overall producer's and user's accuracy was computed for the three strata/classes (cloud, shadow, and cloud-free pixels).

## 3. Results

### 3.1. Cloud and shadow masks

Figure 5 illustrates the process of cloud and shadow detection by showing some intermediate and final results from Landsat scene p42r28. Figures 5(a) and (b) show relaxed and restricted cloud and shadow masks. The relaxed mask (Figure 5(a)) has a large number

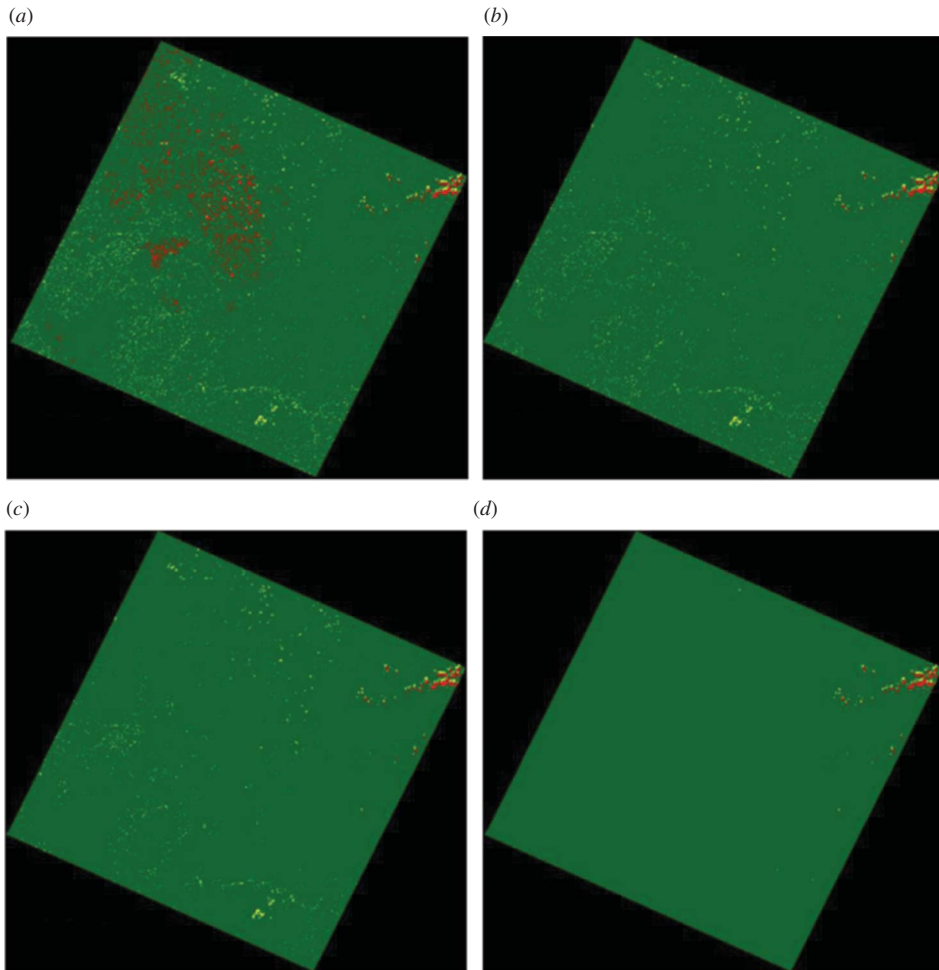


Figure 5. Cloud (displayed as red) and shadow (displayed as yellow) masks of p42r48: (a) relaxed mask; (b) restricted mask; (c) integrated mask; (d) constrained final mask.

of cloud commission errors from agricultural lands in the upper left of the scene. The restricted mask (Figure 5(b)) seems to have no cloud commission errors because extra information from thermal bands helped remove many false clouds. Both masks have a large number of shadow commission errors due to the forest pixels that are in terrain shadow and prevalent in the scene. Figure 5(c) shows the integrated cloud and shadow mask from the relaxed and restricted masks. The cloud patches in the integrated mask look similar to the clouds in the restricted mask, yet with the full shape of real cloud patches (upper right area) as in the relaxed mask. Figure 5(d) shows the final constrained cloud and shadow mask after geometric constraints were applied to the integrated mask. The constrained cloud and shadow mask (Figure 5(d)) is apparently much cleaner than the integrated mask (Figure 5(c)) because many false shadow patches were removed.

Table 2 summarizes the histogram of cloud and shadow masks for all five test scenes. As expected, the restricted cloud and shadow masks have less cloud and shadow than the relaxed masks. In most cases, the pixel numbers of each class in the integrated mask

Table 2. Pixel numbers of cloud and shadow classes of different cloud and shadow masks.

Scene	Class	Cloud and shadow masks					
		Relaxed	Restricted	Restricted*	Integrated	Integrated (%)	Constrained
p43r32	Cloud	4,389,150	3,898,768	88.8	4,270,145	97.3	4,270,145
	Shadow	2,282,842	1,909,689	83.7	2,149,952	94.2	2,128,016
p42r28	Cloud	685,435	67,649	9.9	67,370	9.8	67,370
	Shadow	590,140	412,202	69.8	218,325	37.0	41,099
p33r33	Cloud	826,276	735,682	89.0	806,857	97.6	806,857
	Shadow	1,086,046	576,993	53.1	429,051	39.5	90,613
p18r34	Cloud	1,982,351	123,219	6.2	397,624	20.1	397,624
	Shadow	5,277,234	4,257,175	80.7	4,590,042	87.0	51,199
p12r30	Cloud	1,886,035	1,791,496	95.0	1,860,469	98.6	1,860,469
	Shadow	238,152	115,383	48.4	196,640	82.6	177,155

Note: \*Pixel numbers and percentages relative to relaxed mask were calculated for restricted, integrated, and constrained masks.

are between the relaxed and restricted masks. Owing to a cleaning process (i.e. eliminating cloud and shadow patches with sizes less than 5 pixels in the restricted mask and 10 pixels in the relaxed mask) embedded in the integration procedure, an integrated mask may have fewer cloud or shadow pixels than the restricted mask. The restricted mask of p43r32 has 88.8% of the cloud or shadow pixels in the restricted mask. The restricted mask of p43r32 has 88.8% of the cloud pixels and 83.7% of the shadow pixels in the relaxed mask, whereas the integrated mask has 97.3% of the cloud pixels and 94.2% of the shadow pixels in the relaxed mask. The thermal band does not have a larger impact on the cloud and shadow detection for p43r32. The restricted mask of p42r28 has only 9.9% of the cloud pixels and 69.8% of the shadow pixels in the relaxed mask, and the integrated mask has a similar number of cloud pixels, but about half the number of shadow pixels in the restricted mask. The restricted mask of p18r34 has 6.2% of the cloud pixels and 80.7% of the shadow pixels, and the integrated mask has 20.1% of the cloud pixels and 7.0% of the shadow pixels in the relaxed mask. Integration was very effective for p42r28 and p18r34 in removing many of the false clouds (some agricultural lands in p42r28 and some mining areas in p18r34) in the relaxed masks, while keeping the full shape of real clouds compared to the restricted masks. The effect of integration was less noticeable on cloud classes in p33r33 and p12r30 because the cloud classes in their relaxed mask have low commission errors; therefore, the integration masks have cloud pixel numbers very close to the relaxed masks. In general, the integration process is less effective in reducing false shadows than in reducing false clouds because the thermal band has less effect on separating false shadows from real shadows than separating false clouds from real clouds.

The constrained cloud and shadow masks reduced false shadow patches dramatically after three geometric constraints were applied to the integrated masks, except for p12r30 (Table 2). The constrained mask of p42r28 has 7.0% of the shadow pixels in the relaxed mask, and the constrained mask of p18r34 has only 1.0% of the shadow pixels in the relaxed mask. The scene of p12r30 does not have much shadow commission error in the first place because the topography is relatively flat and the landscape is mainly composed of forest and water. Figure 6 shows the final constrained cloud and shadow masks for the other four path/rows besides p42r28. Five subset examples of the constrained cloud and shadow mask are shown in Figures 7(c), 8(c), 9(c), 10(c), and 11(c). Clouds were detected very well for four scenes (p42r28, p33r33, p18r34, and p12r28) without visually notable commission and omission errors. Some cloud omission errors were observed at the cloud edges and small patches in p43r32. Omission errors were observed in shadow detection mainly because of confusion with terrain shadows in p42r28, small patch sizes in p33r33, and the forest and water in p12r30. We have relatively high commission errors of shadow in p18r34 because geometric constraints could not remove those errors due to the clustered clouds. Even though some shadow pixels are missed, the majority of the shadow patches are captured.

### 3.2. Accuracy assessment

To assess the detection accuracy quantitatively, random points, which were initially created on the constrained cloud and shadow mask, were applied to other intermediate cloud and shadow masks from the model. Table 3 shows the producer's, user's, and overall accuracy of all the classes (cloud, shadow, and clear) from those masks. Integrated cloud and shadow masks generally have the highest overall accuracy compared to relaxed and restricted masks. The scene of p12r30 is an exception because even the relaxed cloud and shadow mask has few commission errors with a user's accuracy of 100% and 97.8% for cloud and shadow, respectively. Additionally, the overall accuracies for the integrated (95.4%) and relaxed (97.4%) masks are very close.



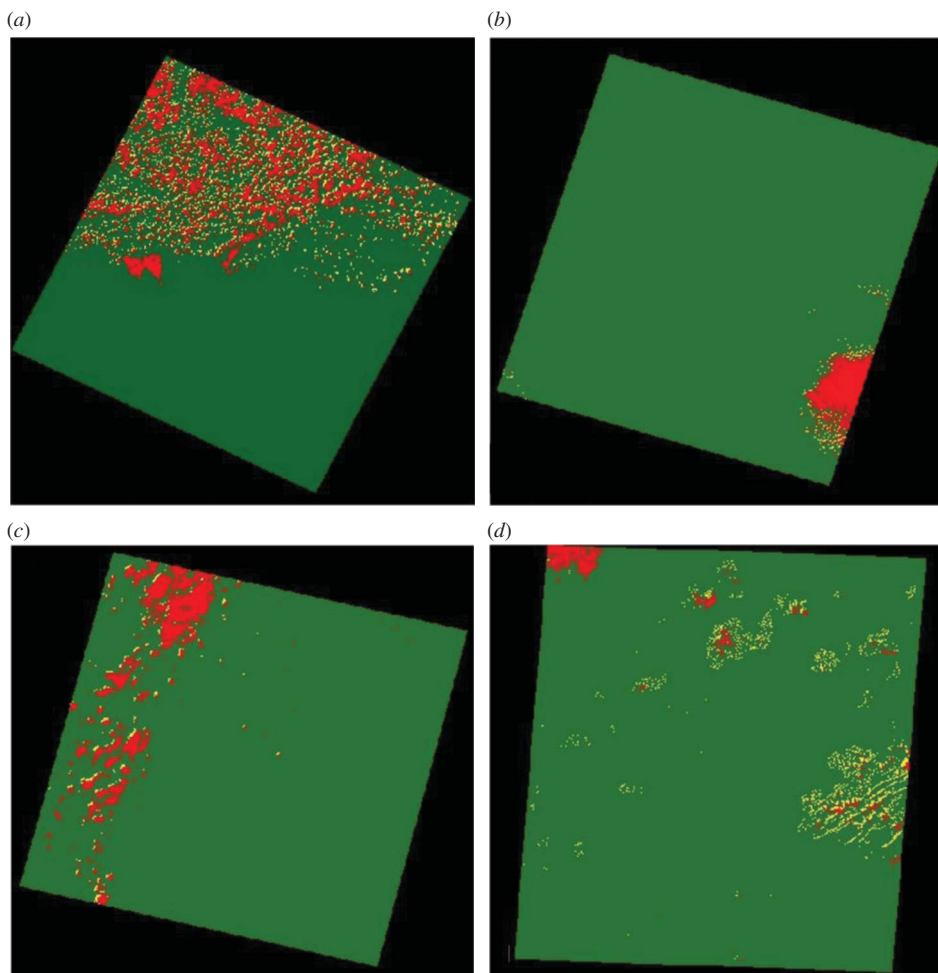


Figure 6. The constrained cloud (displayed as red) and shadow (displayed as yellow) masks for the other four Landsat path/rows: (a) p43r32; (b) p33r33; (c) p12r28; (d) p18r34.

The constrained cloud and shadow masks clearly have equal or higher overall accuracy than the integrated masks because the model of geometric constraints conservatively removed large amounts of false shadow (Table 3). The lowest producer's accuracy of cloud from the constrained masks was 93.9%, which was in the scene of p43r32. Those omission errors were mainly caused by cloud edges. The lowest user's accuracy of cloud from the constrained masks was 89.0%, which was in the scene of p18r34 where some mining areas were mistakenly classified as clouds. The scene of p18r34 also has the worst user's accuracy of 33.7% for shadow because of the mountain terrain; however, it still has a producer's accuracy of 97.1% for shadow.

### 3.3. Cloud-filling effects

Figures 7(d), 8(d), 9(d), 10(d), and 11(d) show the images after the cloud and shadow filling. Visually, no stripes or spatial shifts or boundaries on any filled image exist. Some dark areas still exist in the filled image of p42r28 (Figure 7(d)) because of the omission

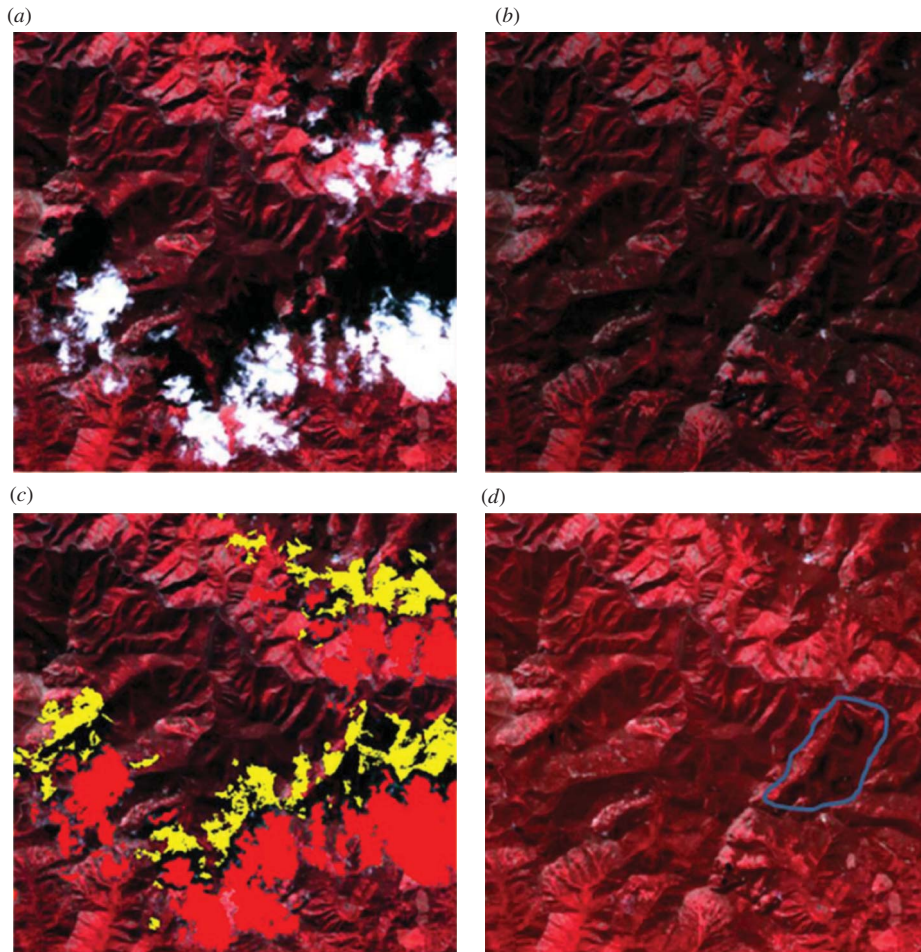


Figure 7. Landsat scene p42r28: (a) target image (11 September 1999); (b) reference image (28 August 2000); (c) cloud (displayed as red) and shadow (displayed as yellow) mask on the target image; (d) cloud-filled image (black areas within the blue polygon are shadows that were not captured by cloud and shadow detection).

errors in shadow detection from the buffered cloud and shadow mask. The areas filled according to the algorithm blended in with the rest of the areas well, and one cannot discern where the original clouds and shadows were located without looking at the target image. Despite the different phenology shown in between the target image and the reference image, the filled image of p18r34 (Figure 9(d)) still appears seamless; this is because we filled the image using the pixels from the target image itself and only used the reference image for identifying the locations of those similar pixels. The spatial and spectral integrity of linear features in the filled images, such as roads and rivers, is maintained as well.

#### 4. Discussion

Integrated cloud and shadow masks generally have the highest overall accuracy compared to relaxed and restricted masks, indicating that the integration process can intelligently

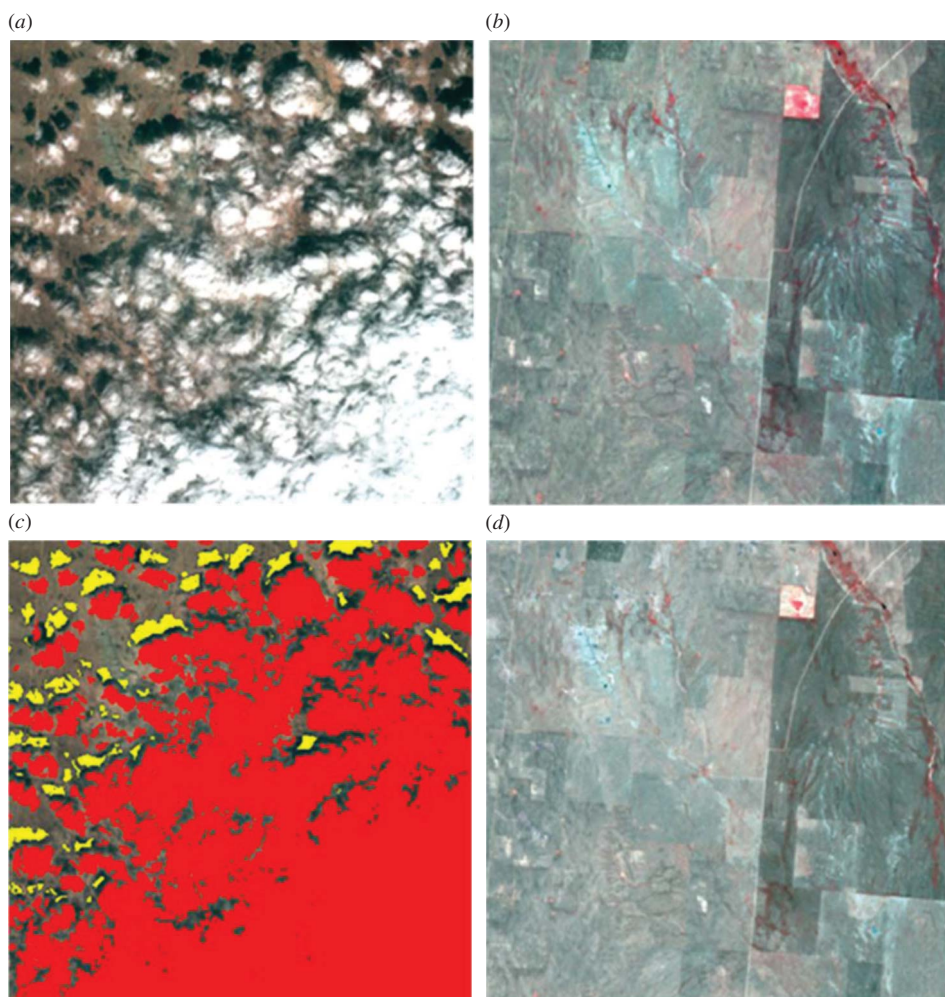


Figure 8. Landsat scene p33r33: (a) target image (16 October 2000); (b) reference image (14 September 2000); (c) cloud (displayed as red) and shadow (displayed as yellow) mask on the target image; (d) cloud-filled image.

combine the advantages from both relaxed and restricted cloud and shadow masks. Integrated cloud and shadow masks basically keep the full shape of real cloud and shadow patches, while enabling the removal of false cloud and shadow patches. The process of integration is conceptually similar to a multi-scale object classification. The integration method is a useful way to integrate the spectral information from visible, infrared, and thermal bands by combining relaxed and restricted masks. The integration mask may lose some small patches of cloud and shadow compared to the relaxed mask when the restricted mask does not detect any cloud or shadow pixel in those small patches.

The final cloud and shadow masks (i.e. the constrained masks) were shown to have high accuracies, especially for cloud. The lowest producer's accuracy of cloud is 93.9% and the lowest user's accuracy is 89.0% among the five test scenes. The low cloud omission error from our model is likely due to the sensitivity of the blue bands, and the low to moderate



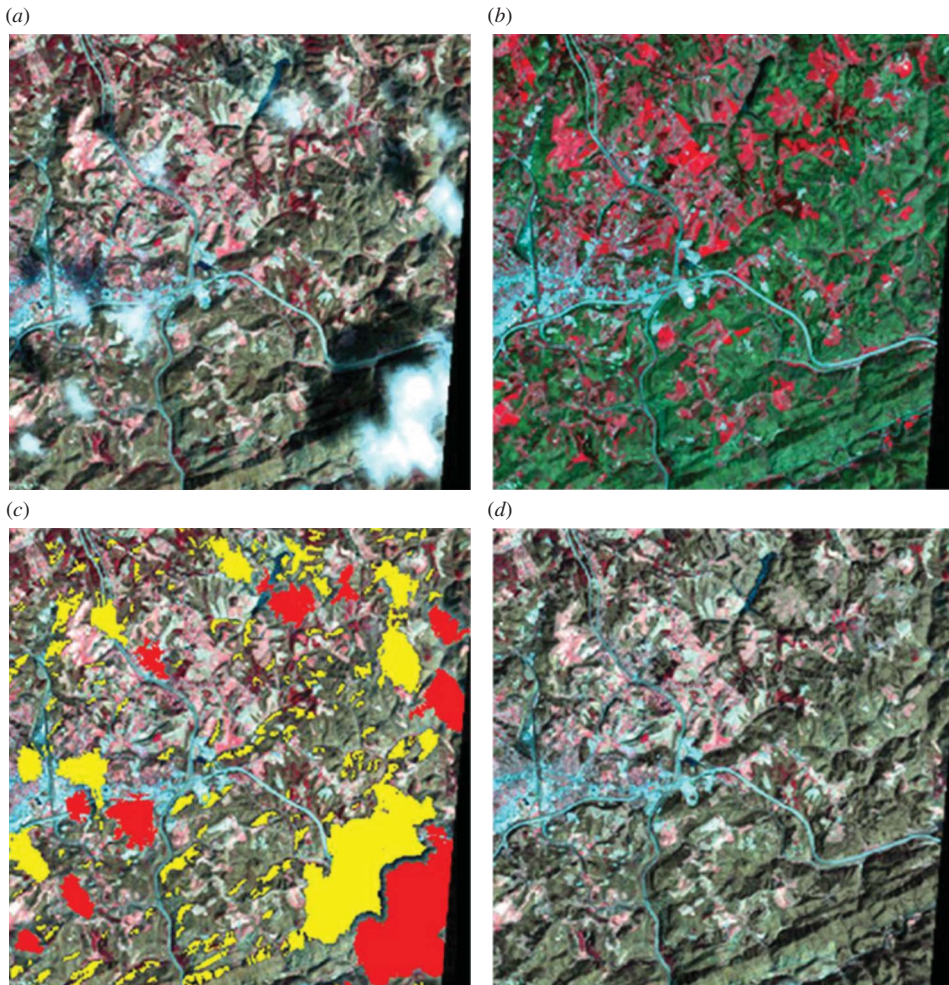


Figure 9. Landsat scene p18r34: (a) target image (5 March 2000); (b) reference image (30 April 2000); (c) cloud (displayed as red) and shadow (displayed as yellow) mask on the target image; (d) cloud-filled image.

commission error can be attributed to the flexible use of the thermal and temporal information from two-date Landsat images. Shadows are difficult to map, and geometric constraints are shown to help reduce the false shadow pixels tremendously. Shadow pixel numbers in the constrained masks cover a range from 1.0% to 86.3% of the relaxed masks. The worst scene, p18r34, had a producer's accuracy of 97.1% and a user's accuracy of 33.7% because some false shadows trapped in cloud clusters could not be removed according to the three geometric constraints we applied. Individual users can adjust the threshold values in the first step of detecting cloud and shadow and some parameters in the last step of removing the false shadows to achieve better results for their scenes.

All five filled images are visually seamless regardless of their phenology differences for the target and reference images and the cloud and shadow patch size. The linear features on those filled images retain their spatial and spectral continuity. Our new method is also easy to implement and less time-consuming. We searched the SSG using the entire scene

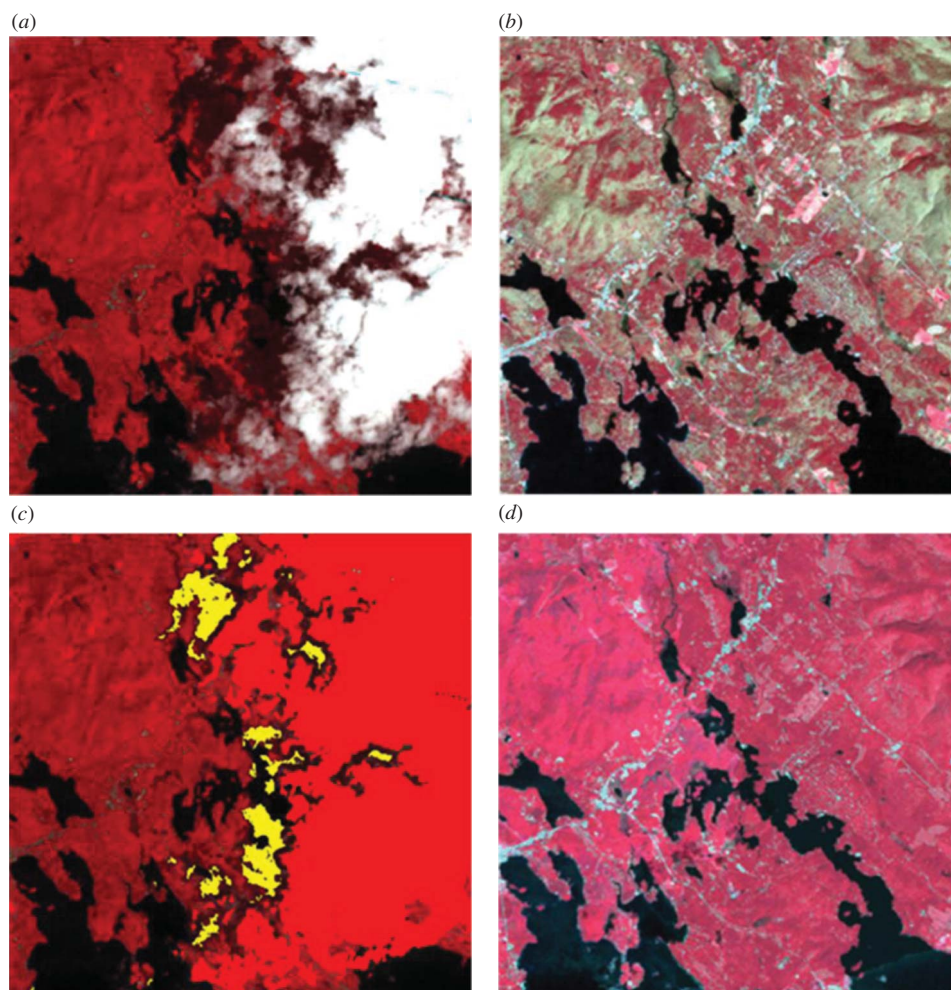


Figure 10. Landsat scene p12r30: (a) target image (16 June 2006); (b) reference image (29 April 2006); (c) cloud (displayed as red) and shadow (displayed as yellow) mask on the target image; (d) cloud-filled image.

for simplicity and efficiency. However, if those pixels within one SSG do not change in a systematic way between the target and reference image dates (i.e. SSG pixels from the reference image do not belong to or deviate far from those of an SSG on the target image), the model will not work well for some pixels. One improvement for future research is to restrict the spatial distance as well as the spectral similarity.

## 5. Conclusion

As Landsat data became freely available, an increased desire to use more of the archive for multiple terrestrial applications, including data sets with clouds, arose. The need to automatically detect cloud and shadow and restore those areas with clean pixel values has become compelling. Our method of cloud and shadow detection flexibly makes full use of visible, SWIR, and thermal spectral information, as well as temporal information



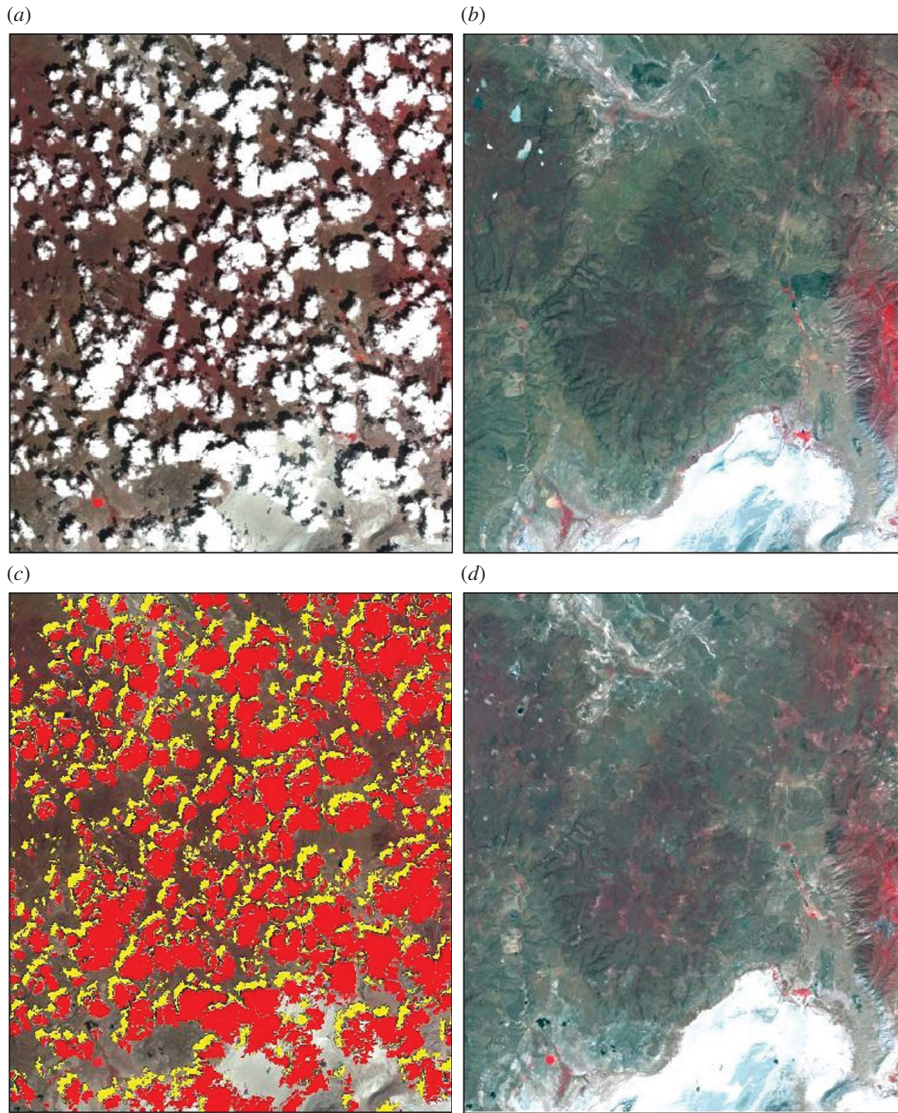


Figure 11. Landsat scene p43r32: (a) target image (20 June 2010); (b) reference image (11 July 2006); (c) cloud (displayed as red) and shadow (displayed as yellow) mask on the target image; (d) cloud-filled image.

and geometric relationships between clouds and their shadows. The test results from this study illustrate the wide applicability of the cloud and shadow detection for Landsat scenes selected for the NLCD production. The model can reduce omission error effectively without introducing large commission error. Besides the cloud and shadow detection method, we introduced a new concept of SSG and the model developed based on the concept was shown to be effective in replacing the contaminated area with reasonable values. Our methods do not require that the reference image be acquired under similar illumination conditions and phenology as the target image.



Table 3. Accuracy of cloud and shadow masks for the five Landsat test scenes.

Scene	Version	Class	Accuracy (%)		
			Producers	Users	Overall
p43r32	Constrained	Cloud	93.9	93.9	95.3
		Shadow	92.9	96.3	
		Clear	96.7	95.6	
	Integrated	Cloud	93.9	93.9	95.3
		Shadow	92.9	96.3	
		Clear	96.7	95.6	
	Relaxed	Cloud	93.9	92.4	95.3
		Shadow	94.6	96.4	
		Clear	96.1	96.1	
	Restricted	Cloud	93.9	100.0	95.7
		Shadow	87.5	96.1	
		Clear	98.9	94.2	
p42r28	Constrained	Cloud	100.0	100.0	99.3
		Shadow	100.0	96.0	
		Clear	99.0	100.0	
	Integrated	Cloud	100.0	100.0	99.0
		Shadow	100.0	94.1	
		Clear	98.5	100.0	
	Relaxed	Cloud	100.0	94.3	98.0
		Shadow	100.0	94.1	
		Clear	97.0	100.0	
	Restricted	Cloud	98.0	100.0	98.7
		Shadow	100.0	94.1	
		Clear	98.5	99.5	
p33r33	Constrained	Cloud	100.0	100.0	96.3
		Shadow	85.4	87.5	
		Clear	97.6	97.1	
	Integrated	Cloud	100.0	100.0	96.0
		Shadow	85.4	85.4	
		Clear	97.1	97.1	
	Relaxed	Cloud	100.0	100.0	96.0
		Shadow	97.6	78.4	
		Clear	94.7	99.5	
	Restricted	Cloud	94.3	100.0	93.0
		Shadow	65.9	87.1	
		Clear	98.1	92.2	
p18r34	Constrained	Cloud	97.8	89.0	83.6
		Shadow	97.1	33.7	
		Clear	79.1	98.7	
	Integrated	Cloud	97.8	89.0	75.6
		Shadow	97.1	24.1	
		Clear	68.4	98.5	
	Relaxed	Cloud	97.8	81.7	73.4
		Shadow	97.1	23.8	
		Clear	65.4	98.4	
	Restricted	Cloud	37.4	97.1	69.8
		Shadow	85.7	25.6	
		Clear	76.4	81.9	
p12r30	Constrained	Cloud	98.2	100.0	95.4
		Shadow	82.2	97.8	
		Clear	99.3	93.0	

(Continued)

Table 3. (Continued).

Scene	Version	Class	Accuracy (%)		
			Producers	Users	Overall
	Integrated	Cloud	98.2	100.0	95.4
		Shadow	82.2	97.8	
		Clear	99.3	93.0	
	Relaxed	Cloud	98.2	100.0	97.4
		Shadow	91.6	98.0	
		Clear	99.3	96.2	
	Restricted	Cloud	90.1	100.0	85.2
		Shadow	41.1	100.0	
		Clear	100.0	79.2	

Both cloud and shadow detection and filling models were developed in ERDAS IMAGINE (Leica Geosystems, Canton St. Gallen, Switzerland), and each model required target and reference image inputs. As far as the limitations of the proposed method are concerned, it requires the reference image to have no overlapping clouds with the target scene and few land-cover changes between the filling image pair. In some cases, some confusion between cloud shadows and water and terrain shadows may occur. Even with these limitations, our approach, which is potentially appropriate for preparing images for the NLCD production, is simple, practical, and efficient, and it provides reasonably good results.

### Acknowledgements

The authors thank Clayton Kingdon and Peter Wolter for their valuable suggestions and discussions during the method development. We thank Thomas Adamson and Craig Walters for their great editing. We also thank James Vogelmann, Shengli Huang, and Lei Ji for their insightful reviews and comments to the article. This work was performed under US Geological Survey (USGS) contract numbers G08PC91508 and G10PC00044. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the US government.

### References

- Bréon, F., and S. Colzy. 1999. "Cloud Detection from the Spaceborne POLDER Instrument and Validation against Surface Synoptic Observations." *American Meteorological Society* 38: 777–85.
- Chander, G., C. Huang, L. Yang, C. Homer, and C. Larson. 2009. "Developing Consistent Landsat Data Sets for Large Area Applications: The MRLC 2001 Protocol." *IEEE Geoscience and Remote Sensing Letters* 6: 777–81.
- Chen, J., X. Zhu, J. E. Vogelmann, F. Gao, and S. Jin. 2011. "A Simple and Effective Method for Filling Gaps in Landsat ETM+ SLC-Off Images." *Remote Sensing of Environment* 115: 1053–64.
- Choi, H., and R. Bindshadler. 2004. "Cloud Detection in Landsat Imagery of Ice Sheets Using Shadow Matching Technique and Automatic Normalized Difference Snow Index Threshold Value Decision." *Remote Sensing of Environment* 91: 237–42.
- Goward, S., J. Irons, S. Franks, T. Arvidson, D. Williams, and J. Faundeen. 2006. "Historical Record of Landsat Global Coverage: Mission Operations, NSLRSDA, and International Cooperator Stations." *Photogrammetric Engineering & Remote Sensing* 72: 1155–69.
- Hagolle, O., M. Huc, V. Pascual, and G. Dedieu. 2010. "A Multi-Temporal Method for Cloud Detection, Applied to FORMOSAT-2, VENµS, LANDSAT and SENTINEL-2 Images." *Remote Sensing of Environment* 114: 1747–55.
- Helmer, E. H., and B. Ruefenacht. 2005. "Cloud-Free Satellite Image Mosaics with Regression Trees and Histogram Matching." *Photogrammetric Engineering & Remote Sensing* 71: 1079–89.

- Herold, N. D., D. Cunningham, and B. K. Wylie. 2003. "Image Processing for the Removal and Data Generation of Cloud Covered Areas through Classification and Regression Tree (CART) Analysis." Paper presented at ASPRS 2003 annual conference proceedings, Anchorage, AK, May 5–9.
- Ho, A. T. S., and Z. Cai. 1996. "Cloud Detection in Satellite Images for Tropical Regions." *Proceedings of SPIE* 2818: 167–71.
- Homer, C., C. Huang, L. Yang, B. K. Wylie, and M. Coan. 2004. "Development of a 2001 National Land-Cover Database for the United States." *Photogrammetric Engineering & Remote Sensing* 70, no. 7: 829–40.
- Huang, C., Nancy, T., S. N. Goward, J. G. Masek, Z. Zhu, J. R. G. Townshend, and J. E. Vogelmann. 2010. "Automated Masking of Cloud and Cloud Shadow for Forest Change Analysis Using Landsat Images." *International Journal of Remote Sensing* 31, no. 20: 5449–64.
- Irish, R. 2000. "Landsat 7 Automatic Cloud Cover Assessment." *Proceedings of SPIE* 4049: 348–55.
- Irish, R. R., J. L. Barker, S. N. Goward, and T. J. Arvidson. 2006. "Characterization of the Landsat-7 ETM Automated Cloud-Cover Assessment (ACCA) Algorithm." *Photogrammetric Engineering & Remote Sensing* 72: 1179–88.
- Ju, J., and D. P. Roy. 2008. "The Availability of Cloud-Free Landsat ETM+ Data over the Conterminous United States and Globally." *Remote Sensing of Environment* 112: 1196–211.
- Le Hégarat-Masclé, S., and C. André. 2009. "Use of Markov Random Fields for Automatic Cloud/Shadow Detection on High Resolution Optical Images." *Photogrammetry and Remote Sensing* 64: 351–66.
- Martinuzzi, S., W. A. Gould, and O. M. Ramos González. 2007. *Creating Cloud-Free Landsat ETM+ Data Sets in Tropical Landscapes: Cloud and Cloud-Shadow Removal*, General Technical Report IITF-GTR-32. International Institute of Tropical Forestry (IITF), U.S. Department of Agriculture (USDA), Forest Service. Puerto Rico: IITF.
- Roy, D. P. 2000. "The Impact of Misregistration upon Composited Wide Field of View Satellite Data and Implications for Change Detection." *IEEE Transactions on Geoscience and Remote Sensing* 38: 2017–32.
- Roy, D. P., J. Ju, K. Kline, P. L. Scaramuzza, V. Kovalsky, M. Hansen, T. R. Loveland, E. Vermote, and C. Zhang. 2010. "Web-Enabled Landsat Data (WELD): Landsat ETM+ Composited Mosaics of the Conterminous United States." *Remote Sensing of Environment* 114: 35–49.
- Roy, D. P., J. Ju, P. Lewis, C. Schaaf, F. Gao, M. Hansen, and E. Lindquist. 2008. "Multi-Temporal MODIS–Landsat Data Fusion for Relative Radiometric Normalization, Gap Filling, and Prediction of Landsat Data." *Remote Sensing of Environment* 112: 3112–30.
- Vogelmann, J. E., S. M. Howard, L. Yang, C. R. Larson, B. K. Wylie, and J. N. Van Driel. 2001. "Completion of the 1990's NLCD for the Conterminous US." *Photogrammetric Engineering & Remote Sensing* 67: 650–62.
- Wang, B., A. Ono, K. Muramatsu, and N. Fujiwara. 1999. "Automated Detection and Removal of Clouds and their Shadows from Landsat TM Images." *IEEE Transactions on Information and Systems* 2: 453–60.
- Williams, D. L., S. Goward, and T. Arvidson. 2006. "Landsat: Yesterday, Today, and Tomorrow." *Photogrammetric Engineering & Remote Sensing* 72: 1171–8.
- Woodcock, C. E., R. Allen, M. Anderson, A. Belward, R. Bindschadler, W. Cohen, F. Gao, S. N. Goward, D. Helder, E. Helmer, R. Nemani, L. Oreopoulos, J. Schott, P. S. Thenkabail, E. F. Vermote, J. Vogelmann, M. A. Wulder, and R. Wynne. 2008. "Free Access to Landsat Imagery." *Science* 320: 1011.
- Zhang, C., W. Li, and D. Travis. 2007. "Gaps-Fill of SLC-off Landsat ETM Plus Satellite Image Using a Geostatistical Approach." *International Journal of Remote Sensing* 28: 5103–22.